Locally recoverable codes on algebraic curves

Alexander Barg[a] Itzhak Tamo[b] Serge Vlăduţ[c]

***Abstract*—A code over a finite alphabet is called locally recoverable (LRC code) if every symbol in the encoding is a function of a small number (at most r) other symbols. A family of linear LRC codes that generalize the classic construction of Reed-Solomon codes was constructed in a recent paper by I. Tamo and A. Barg (*IEEE Trans. Inform. Theory*, vol. 60, no. 8, 2014, pp. 4661-4676). In this paper we extend this construction to codes on algebraic curves. We give a general construction of LRC codes on curves and compute some examples, including asymptotically good families of codes derived from the Garcia-Stichtenoth towers. The local recovery procedure is performed by polynomial interpolation over r coordinates of the codevector. We also obtain a family of Hermitian codes with two disjoint recovering sets for every symbol of the codeword.**

I. Introduction: LRC Codes

The notion of LRC codes is motivated by applications of coding to increasing reliability and efficiency of distributed storage systems. Following [3], we say that a code $\mathcal{C} \subset \mathbb{F}_q^n$ has locality r if every symbol of the codeword $x = (x_1, ..., x_n) \in \mathcal{C}$ can be recovered from a subset of r other symbols of x (i.e., is a function of some other r symbols $x_{i_1}, x_{i_2}, \ldots, x_{i_r}$). In other words, this means that for every $i \in [n]$ there exists a subset of coordinates $I_i \subset [n]\backslash i, |I_i| \le r$ such that the value of x_i is be found from the restriction of $\mathcal{C}$ to the coordinates in I_i. The subset I_i is called a *recovering set* for the ith coordinate of the codeword.

The formal definition is as follows. Given $a \in \mathbb{F}_q$ consider the sets of codewords

$$\mathcal{C}(i, a) = \{x \in \mathcal{C} : x_i = a\}, \quad i \in [n].$$

The code $\mathcal{C}$ is said to have *all-symbol locality* r if for every $i \in [n]$ there exists a subset $I_i \subset [n]\backslash i, |I_i| \le r$ such that the restrictions of the sets $\mathcal{C}(i, a)$ to the coordinates in I_i for different a are disjoint:

$$\mathcal{C}_{I_i}(i, a) \cap \mathcal{C}_{I_i}(i, a') = \emptyset, \quad a \ne a'. \tag{1}$$

We use the notation (n, k, r) to refer to the parameters of the code, where $k = \log_q |C|$.

This definition can be extended to codes with *multiple recovering sets*; see, e.g., [5], [8]. For instance, suppose that for every symbol $i \in [n]$ condition (1) holds true for the subset $I_i, |I_i| = r_1$ as well as for some other subset $J_i \subset [n], |J_i| = r_2$. In this case we say that every symbol has two recovering sets, and use the notation $(n, k, \{r_1, r_2\})$ for the code parameters. In the literature it is also often assumed that $I_i \cap J_i = \emptyset, i \in [n]$, and we include this in the definition of LRC codes with two recovering sets (LRC(2) codes) used in this paper.

Let $\mathcal{C}$ be an (n, k, r) LRC code of cardinality q^k. The minimum distance of $\mathcal{C}$ is known to satisfy [3], [4]

$$d \le n - k - \left\lceil \frac{k}{r} \right\rceil + 2. \tag{2}$$

The bound (2) is an extension of the classical Singleton bound of coding theory, which is attained by the well-known family of Reed-Solomon (RS) codes. RS-like codes with the LRC property whose parameters meet the bound (2) were recently constructed in [9]. Unlike some other known constructions, e.g., [7], [11], the codes [9] are constructed over finite fields of cardinality comparable to the code length n. The cyclic case of the construction in [9] is studied in the recent paper [10].

Classical RS codes can be viewed as a special case of the general construction of geometric Goppa codes; in particular, good codes are obtained from families of curves with many rational points [12]. Motivated by this approach, in this paper we take a similar view of the construction of the evaluation codes of [9]. Interpreting these codes as codes on algebraic curves, we present a general construction of algebraic geometric LRC codes and compute the parameters of codes for some particular choices of curves. Similarly to [9], local recovery can be performed by interpolating a univariate polynomial over r coordinates of the codeword. The RS-like codes in [9] can be extended to multiple recovering sets, and here we point out one such extension in the case of Hermitian codes.

II. The construction of [9]

Let us briefly recall the construction of [9]. Our aim is to construct an LRC code over $\mathbb{F}_q$ with the parameters (n, k, r), where $n \le q$. We additionally assume that $(r + 1)|n$ and $r|k$, although both the constraints can be lifted by adjustments to the construction presented below [9]. Let $g(x) \in \mathbb{F}_q[x]$ be a polynomial of degree $r + 1$ such that there exists a partition $\mathcal{A} = \{A_1, \ldots, A_{\frac{n}{r+1}}\}$ of a set of points $A = \{P_1, \ldots, P_n\} \subset \mathbb{F}_q$ into subsets of size $r + 1$ such that g is constant on each set $A_i \in \mathcal{A}$.

Consider a k-dimensional linear subspace $V \subset \mathbb{F}_q[x]$ generated by the set of polynomials

$$(g(x)^j x^i,\ i = 0, \ldots, r - 1; j = 0, \ldots, \frac{k}{r} - 1). \tag{3}$$

Given $a = (a_{ij}, i = 0, \ldots, r - 1; j = 0, \ldots, \frac{k}{r} - 1) \in \mathbb{F}_q^k$ let

$$f_a(x) = \sum_{i=0}^{r-1} f_i(x) x^i, \text{ where } f_i(x) = \sum_{j=0}^{\frac{k}{r}-1} a_{ij} g(x)^j, \quad i = 0, \ldots, r - 1. \tag{4}$$

[a] Dept. of ECE and ISR, University of Maryland, College Park, MD 20742 and IITP, Russian Academy of Sciences, Moscow, Russia. Email abarg@umd.edu. Research supported by NSF grants CCF1422955, CCF1217894, and CCF1217245.

[b] Dept. of EE-Systems, Tel Aviv University, Tel Aviv, Israel. Research done while at the Institute for Systems Research, University of Maryland, College Park, MD 20742. Email zactamo@gmail.com. Research supported by NSF grant CCF1217894.

[c] Institut de Mathématiques de Marseille, Aix-Marseille Université, IML, Luminy case 907, 13288 Marseille, France, and IITP, Russian Academy of Sciences, Moscow, Russia. Email serge.vladuts@univ-amu.fr.

Now define the code $\mathcal{C}$ as the image of the linear evaluation map

$$\begin{aligned} e : V &\to \mathbb{F}_q^n \\ f_a &\mapsto (f_a(P_i), i = 1, \dots, n). \end{aligned} \tag{5}$$

As shown in [9], $\mathcal{C}$ is an (n, k, r) LRC code whose minimum distance d meets the bound (2) with equality.

To construct examples of codes using this approach we need to find polynomials and partitions of points of the field that satisfy the above assumptions. As shown in [9], one can take $g(x) = \prod_{\beta \in H}(x - \beta)$, where H is a subgroup of the additive or the multiplicative group of $\mathbb{F}_q$ (see also the example in the next section). In this case $r = |H| - 1$, and the corresponding set of points A can be taken to be any collection of the cosets of the subgroup H in the full group of points. In this way we can construct codes of length $n = m(r + 1)$, where $m \geq 1$ is an integer that does not exceed $(q-1)/|H| = (q-1)/(r+1)$ or $q/|H| = q/(r + 1)$ depending on the choice of the group.

III. Algebraic geometric LRC codes

As above, let us fix a finite field $\Bbbk = \mathbb{F}_q, q = p^a$ of characteristic p. To motivate our construction, consider the following example.

Example 1: Let H be a cyclic subgroup of $\mathbb{F}_{13}^*$ generated by 3 and let $g(x) = x^3$. Let $r = 2, n = 9, k = 4$, and choose $A = \{1, 2, 3, 4, 5, 6, 9, 10, 12\}$. We obtain $\mathcal{A} = \{A_1, A_2.A_3\}$, where

$$\begin{aligned} A_1 = \{1, 3, 9\}, \quad A_2 = \{2, 6, 5\}, \quad A_3 = \{4, 12, 10\} \\ g(A_1) = 1 \qquad g(A_2) = 8 \qquad g(A_3) = 12 \end{aligned} \tag{6}$$

The set of polynomials (3) has the form $(1, x, x^3, x^4)$. In this case Construction (5) yields a $(9, 4, 2)$ LRC code with distance $d = 5$ [9].

This construction can be given the following geometric interpretation: the polynomial g defines a mapping $g : \mathbb{P}^1 \to \mathbb{P}^1$ of degree $r + 1 = 3$ such that the preimage of every point in $g(A)$ consists of "rational" points (i.e., $\mathbb{F}_q$-points). This suggests a generalization of the construction to algebraic curves which we proceed to describe (note Example 2 on the following page that may make it easier to understand the general case).

Let X and Y be smooth projective absolutely irreducible curves over $\Bbbk$. Let $g : X \to Y$ be a rational separable map of curves of degree $r + 1$. As usual, denote by $\Bbbk(X)$ ($\Bbbk(Y)$) the field of rational functions on X (resp., Y). Let $g^* : \Bbbk(Y) \to \Bbbk(X)$ be the function that acts on $\Bbbk(Y)$ by $g^*(f)(P) = f(g(P))$, where $f \in \Bbbk(Y), P \in X$. The map g^* defines a field embedding $\Bbbk(Y) \hookrightarrow \Bbbk(X)$, and we identify $\Bbbk(Y)$ with its image $g^*(\Bbbk(Y)) \subset \Bbbk(X)$.

Since g is separable, the primitive element theorem implies that there exists a function $x \in \Bbbk(X)$ such that $\Bbbk(X) = \Bbbk(Y)(x)$, and that satisfies the equation

$$x^{r+1} + b_r x^r + \dots + b_0 = 0,$$

where $b_i \in \Bbbk(Y)$. The function x can be considered as a map $x : X \to \mathbb{P}^1_{\Bbbk}$, and we denote its degree $\deg(x)$ by h.

Example 1: (continued) For instance, in the above example, we have $X = \mathbb{P}^1, Y = \mathbb{P}^1$, and the mapping g is given by $y = x^3$. We obtain $\Bbbk(Y) = \Bbbk(x^3) = \Bbbk(y)$, $\Bbbk(X) = \Bbbk(y)(x)$, where x satisfies the equation $x^3 - y = 0$. Note that in this case $b_r = b_{r-1} = ... = b_1 = 0, b_0 = y$.

The codes that we construct belong to the class of evaluation codes. Let $S = \{P_1, \dots, P_s\} \subset Y(\Bbbk)$ be a subset of $\mathbb{F}_q$-rational points of Y and let Q_∞ be a positive divisor of degree $\ell \geq 1$ whose support is disjoint from S. For instance, one can assume that $Q_\infty = \pi^{-1}(\infty)$ for a projection $\pi : Y \to \mathbb{P}^1_{\Bbbk}$. To construct our codes we introduce the following set of fundamental assumptions with respect to S and g:

$$\begin{aligned} A := g^{-1}(S) = \{P_{ij}, i = 0, \dots, r, j = 1, \dots, s\} \subseteq X(\Bbbk); \\ g(P_{ij}) = P_j \text{ for all } i, j; \\ b_i \in L(n_i Q_\infty), \quad i = 0, 1, \dots, r, \end{aligned} \tag{7}$$

for some natural numbers n_i.

Now let $D = tQ_\infty$ be a positive divisor and let $\{f_1, \dots, f_m\}$ be a basis of the linear space $L(D)$. The functions $f_i, i = 1, \dots, m$ are contained in $\Bbbk(Y)$ and therefore are constant on the fibers of the map g. The Riemann-Roch theorem implies that $m \geq t\ell - g_Y + 1$, where g_Y is the genus of Y. Below we assume that $m = t\ell - g_Y + 1$. Consider the $\Bbbk$-subspace V of $\Bbbk(X)$ of dimension rm generated by the functions

$$\{f_j x^i, i = 0, \dots, r - 1, j = 1, \dots, m\} \tag{8}$$

(note an analogy with (3)). Since Q_∞ is disjoint from S, the evaluation map

$$\begin{aligned} e := ev_A : V &\longrightarrow \Bbbk^{(r+1)s} \\ F &\mapsto (F(P_{ij}), i = 0, \dots, r, j = 1, \dots, s) \end{aligned} \tag{9}$$

is well-defined. The image of this mapping is a linear subspace of $\mathbb{F}_q^{(r+1)s}$ (i.e., a code), which we denote by $\mathcal{C}(D, g)$. The code coordinates are naturally partitioned into s subsets $A_j = \{P_{ij}, i = 0, ..., r\}, j = 1, \dots, s$ of size $r + 1$ each; see (7). Assume throughout that, for any fixed j, x takes different values at the points in the set $(P_{ij}, i = 0, \dots, r)$.

Theorem 3.1: The subspace $\mathcal{C}(D, g) \subset \mathbb{F}_q$ forms an (n, k, r) linear LRC code with the parameters

$$\left.\begin{aligned} n &= (r + 1)s \\ k &= rm \geq r(t\ell - g_Y + 1) \\ d &\geq n - t\ell(r + 1) - (r - 1)h \end{aligned}\right\} \tag{10}$$

provided that the right-hand side of the inequality for d a positive integer. Local recovery of an erased symbol $c_{ij} = F(P_{ij})$ can be performed by polynomial interpolation through the points of the recovering set A_j.

Proof: The first relation in (10) follows by construction. The inequality for the distance is also immediate: the function $f_j x^i, f_j \in L(D)$, evaluated on A, can have at most $t\ell(r + 1) + (r - 1)\deg(x)$ zeros. Since we assume that $d \geq 1$, the mapping ev_T is injective, which implies the claim about the dimension of the code. Finally, the functions f_i are constant on the fibers $(P_{ij}, i = 0, \dots, r - 1)$; therefore on each subset A_j the codeword is obtained as an evaluation of the polynomial of the variable x of degree $\rho \leq r-1$. This representation accounts for the fact that coordinate $c_P, P \in A_j$ of the codeword can be found by interpolating a polynomial of degree at most r_1 through the remaining points of A_j. ■

IV. Some code families

Let us give some examples of code families arising from our construction.

A. LRC codes from Hermitian curves

Let $q = q_0^2$, where q_0 is a prime power, let $\Bbbk = \mathbb{F}_q$, and let $X := H$ be the Hermitian curve, i.e., a plane smooth curve of genus $g_0 = q_0(q_0-1)/2$ with the affine equation

$$X : x^{q_0} + x = y^{q_0+1}.$$

The curve X has $q_0^3+1 = q\sqrt{q}+1$ rational points of which one is the infinite point and the remaining q_0^3 are located in the affine plane. There are two slightly different ways of constructing Hermitian LRC codes.

1) Projection on y: Here we construct q-ary $(n, k, r = q_0 - 1)$ LRC codes. Take $Y = \mathbb{P}^1(\Bbbk)$ and take g to be the natural projection defined by $g(x, y) := y$, then the degree of g is $q_0 = r + 1$ and the degree of x is $h = q_0 + 1$. We can write $X(\mathbb{F}_q) = g^{-1}(\mathbb{F}_q) \bigcup Q'_\infty$ where $Q'_\infty \in X$ is the unique point over $\infty \in Y$.

Turning to the code construction, take $S = \Bbbk \subset \mathbb{P}^1, Q_\infty = \infty, \ell = 1$, and $D = tQ'_\infty$ for some $t \geq 1$. We have

$$L(D) = \Big\{ \sum_{i=0}^{t} a_i y^i \Big\} \subset \Bbbk[y].$$

Following the general construction of the previous section, we obtain the following result.

Proposition 4.1: The construction of Theorem 3.1 gives a family of q-ary Hermitian LRC codes with the parameters

$$n = q_0^3, k = (t+1)(q_0-1), r = q_0 - 1$$
$$d \geq n - tq_0 - (q_0-2)(q_0+1).$$

Example 2: Let $q_0 = 3, q = 9, \Bbbk = \mathbb{F}_9$ and consider the Hermitian curve X of genus 3 given by the equation $x^3 + x = y^4$. The curve X has 27 points in the finite plane, shown in Fig.1 below (here $\alpha^2 = \alpha + 1$ in $\mathbb{F}_9$), and one point at infinity.

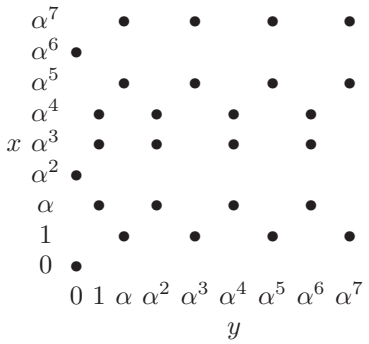

Fig.1: 27 points of the Hermitian curve over $\mathbb{F}_9$.

The columns of the array in Fig. 1 correspond to the fibers of the mapping $g(\cdot, y) = y$, and for every $a \in Y(\mathbb{F}_9) \backslash Q_\infty$ there are 3 points $(\cdot, a) \in X$ lying above it. These triples form the recovering sets $A_1, \ldots, A_9$, similarly to (6). The map $x : X \to \mathbb{P}^1$ has degree $h = 4$. Choosing D in the form $D = tQ'_\infty$ and taking $S = \mathbb{F}_9$ (all the affine points of Y), we obtain an LRC code $\mathcal{C}(D, g)$ with the parameters

$$n = 27, k = 2(t+1), r = 2 \quad (11)$$
$$d \geq 27 - 3t - 4 = 23 - 3t. \quad (12)$$

For instance, take $t = 2$. The basis of functions (8) in this case takes the following form:

$$\{1, y, y^2, x, xy, xy^2\}.$$

To give an example of local decoding, let us compute the codeword for the message vector $(1, \alpha, \alpha^2, \alpha^3, \alpha^4, \alpha^5)$. The polynomial

$$F(x, y) = 1 + \alpha y + \alpha^2 y^2 + \alpha^3 x + \alpha^4 xy + \alpha^5 xy^2$$

evaluates to the codeword

$$\begin{array}{cccccccccc}
\alpha^7 & & & \alpha & & \alpha^7 & & \alpha^5 & & 0 \\
\alpha^6 & \alpha^2 & & & & & & & & \\
\alpha^5 & & & \alpha^6 & & \alpha^4 & & \alpha^2 & & 0 \\
\alpha^4 & & \alpha^7 & & \alpha^3 & & \alpha^5 & & \alpha^5 & \\
x\ \alpha^3 & & \alpha^3 & & \alpha^7 & & \alpha & & \alpha & \\
\alpha^2 & \alpha^3 & & & & & & & & \\
\alpha & & 0 & & 0 & & 0 & & 0 & \\
1 & & & 1 & & \alpha^6 & & \alpha^4 & & 0 \\
0 & 1 & & & & & & & & \\
& 0 & 1 & \alpha & \alpha^2 & \alpha^3 & \alpha^4 & \alpha^5 & \alpha^6 & \alpha^7 \\
& & & & & y & & & &
\end{array}$$

(e.g., $F(0, 0) = 1$, etc.). Suppose that the value at $P = (\alpha, 1)$ is erased. The recovering set for the coordinate P is $\{(\alpha^4, 1), (\alpha^3, 1)\}$, so we compute a linear polynomial $f(x)$ such that $f(\alpha^4) = \alpha^7$ and $f(\alpha^3) = \alpha^3$, i.e., $f(x) = \alpha x - \alpha^2$. Now the coordinate at $(\alpha, 1)$ can be found as $f(\alpha) = 0$. ■

Computing the gap to the Singleton bound (2), we obtain

$$\begin{aligned} d + \frac{k}{r}(r+1) &\geq q_0^3 - tq_0 - (q_0-2)(q_0+1) + q_0(t+1) \\ &= q_0^3 - q_0^2 + 2q_0 + 2 \\ &= n - q + 2\sqrt{q} + 2. \end{aligned} \quad (13)$$

For codes that meet the bound (2) we would have $d + k(r+1)/r = n+2$, so the Singleton gap of the Hermitian LRC codes is at most $q - 2\sqrt{q} = q_0(q_0-2)$. Of course, these codes cannot be Singleton-optimal because their length is much greater than the alphabet size, but the gap in this case is still rather small. For instance in Example 2 we have $d + k(r+1)/r \geq 23 - 3t + 3(t+1) = 26$ while for codes meeting the Singleton bound we would have $d + k(r+1)/r = 29$.

2) Projection on x: Again take $Y = \mathbb{P}^1$ and let $g'(x, y) := x$ be the second natural projection on $\mathbb{P}^1$. There are q_0 points on $\mathbb{P}^1$ that are fully ramified (they have only one point of X above them), namely the points in the set

$$M = \{a \in \mathbb{F}_q : a^{q_0} + a = 0\} \quad (14)$$

(e.g., in Fig. 1 $M = \{0, \alpha^2, \alpha^6\}$). Therefore, every fiber of g' over $\mathbb{F}_q \backslash M$ consists of $\mathbb{F}_q$-rational points since there are in total

$$|\mathbb{F}_q \backslash M| \cdot (q_0 + 1) = q_0^3 - q_0$$

rational points in those fibers. Obviously $|g^{-1}(a) \cap g'^{-1}(b)| \leq 1$ for all $a, b \in \mathbb{F}_q$.

Take $S = \mathbb{F}_q \backslash M$, then $r = q_0$, and clearly $h = \deg(y) = q_0$. We obtain

Proposition 4.2: The construction of Theorem 3.1 gives a family of q-ary Hermitian LRC codes with the parameters

$$n = q_0^3 - q_0, k = (t+1)q_0, r = q_0$$
$$d \geq n - t(q_0+1) - q_0(q_0-1).$$

For instance, in Example 2, taking $t = 2$, we obtain a code of dimension 9 from the basis of functions $\{1, y, y^2, x, xy, xy^2, x^2, x^2y, x^2y^2\}$.

Performing a calculation similar to (13) we obtain the quantity one less than for the first family:

$$d + \frac{k}{r}(r+1) = n - q + 2\sqrt{q} + 1.$$

Remark 4.1: Hermitian LRC codes are in a certain sense optimal for our construction. Note that most known curves with the optimal quotient (number of rational points)/(genus) have the property that for any projection $g : X \to \mathbb{P}^1$ the point $\infty \in \mathbb{P}^1$ is totally ramified (see e.g., the next section). In this case the quantity h satisfies $h \geq n/q$. At the same time, for Hermitian curves, $h = n/q$ (or $(n/q)+1$). Recall also that Hermitian curves are absolutely maximal, i.e. attain the equality in Weil's inequality, and moreover, their genus is maximal for maximal curves.

3) Two recovering sets: The existence of two projections g and g' with mutually transversal fibers suggests that Hermitian LRC codes could be modified, leading to a family of LRC(2) codes with two recovering sets of size $r_1 = q_0 - 1$ and $r_2 = q_0$, respectively. Indeed, let

$$B = g^{-1}(\mathbb{F}_q \backslash \{0\}) = (g')^{-1}(\mathbb{F}_q \backslash M) \subset X/\mathbb{F}_q,$$

$|B| = (q_0^2 - 1)q_0$, where M is defined in (14), and consider the following polynomial space of dimension $(q_0 - 1)q_0$:

$$L := \text{span}\,\{x^i y^j, i = 0, 1, \ldots, r_1 - 1, j = 0, 1, \ldots, r_2 - 1\}.$$

Proposition 4.3: Consider the linear code $\mathcal{C}$ obtained by evaluating the functions in L at the points of B. The code $\mathcal{C}$ has the parameters $(n = (q_0^2 - 1)q_0, k = (q_0 - 1)q_0, \{r_1 = q_0 - 1, r_2 = q_0\})$ and distance

$$d \geq (q_0 + 1)(q_0^2 - 3q_0 + 3). \tag{15}$$

Proof: X is a plane curve of degree $q_0 + 1$, so the Bezout theorem implies that any polynomial of degree $\leq 2q_0 - 3$ has no more than $(q_0 + 1)(2q_0 - 3)$ zeros on X; hence (15). ■

For instance, puncturing the code of Example 2 on the coordinates in M, we obtain an LRC(2) code with the parameters $(24, 6, \{2, 3\})$ and distance $d \geq 12$.

B. Codes from Garcia-Stichtenoth curves

Let $q = q_0^2$ be a square and let $l \geq 2$ be an integer. Define the curve X_l and the functions x_l, z_l inductively as follows:

$$x_0 := 1;\ X_1 := \mathbb{P}^1, \Bbbk(X_1) = \Bbbk(x_1); \tag{16}$$
$$X_l : z_l^{q_0} + z_l = x_{l-1}^{q_0+1}, x_{l-1} := \frac{z_{l-1}}{x_{l-2}} \in \Bbbk(X_{l-1}) \text{ (if } l \geq 3),$$

where $\Bbbk = \mathbb{F}_q$. In particular, $X_2 = H$ is the Hermitian curve. The resulting family of curves is known to be asymptotically maximal [1], [12, p.177], and gives rise to codes with good parameters in the standard error correction problem. Since this family generalizes Hermitian curves, we can expect that it gives rise to two families of codes that extend the constructions of Sect. IV-A. This is indeed the case, as shown below.

1) : To use the general construction that leads to Theorem 3.1 we take the map $g_l : X_l \to X_{l-1}$ to be the natural projection of degree $q_0 = r + 1$. We note that

$$g_l^* : \Bbbk(X_{l-1}) \to \Bbbk(X_l) = \Bbbk(X_{l-1})(z_l). \tag{17}$$

To describe rational points of the curve X_l let $\psi_l : X_l \to \mathbb{P}^1$ be the natural projection of degree q_0^{l-1}, i.e., the map $\psi_l = g_l \circ g_{l-1} \circ \cdots \circ g_2$. Then all the points in the preimage $P_l := \psi_l^{-1}(\mathbb{F}_q^*)$ are $\mathbb{F}_q$-rational, and there are $n_l = q_0^{l-1}(q_0^2 - 1)$ such points. The genus of the curve X_l can be bounded above as

$$G_l \leq q_0^l + q_0^{l-1} = q_0^{l-1}(q_0 + 1) = \frac{n_l}{q_0 - 1}$$

(the exact value of G_l is known [1], but this estimate suffices: in particular, it implies that the curves $X_l, l \to \infty$ are asymptotically maximal). We obtain the following result.

Proposition 4.4: There exists a family of q-ary $(n, k, r = q_0 - 1)$ LRC codes on the curve $X_l, l \geq 2$ with the parameters

$$\left.\begin{aligned} n = n_l &= q_0^{l-1}(q_0^2 - 1) \\ k &\geq r\Big(t - \frac{n_{l-1}}{q_0 - 1} + 1\Big) \\ d &\geq n_l - tq_0 - \frac{2n_l(q_0 - 2)}{q_0^2 - 1} \end{aligned}\right\} \tag{18}$$

where t is any integer such that $G_{l-1} \leq t \leq n_{l-1}$.

Proof: We apply the construction of Theorem 3.1 to $X := X_l, Y := X_{l-1}$, taking the map $g := g_l$, $Q_\infty := P_{\infty,l}$, $D = tQ_\infty$, and $\ell = 1$.

The function x in the general construction in this case is $x = z_l$. To estimate the distance of the code $\mathcal{C}(D, g)$ using (10) we need to find the degree $h = \deg(z_l)$. Toward this end, observe that $z_l = x_l x_{l-1}$, so

$$\deg(z_l) = \deg(x_l) + \deg(x_{l-1}).$$

Let $(x_l)_0^{(l)}$ be the divisor of zeros of x_l on X_l. Recall from [1], Lemma 2.9 that $(x_l)_0^{(l)} = q_0^{l-1} Q_l$, where Q_l is the unique common zero of $x_1, z_2, \ldots, z_l$. Therefore, $\deg(x_l)_0^{(l)} = q_0^{l-1}$ and $\deg(x_{l-1})_0^{(l-1)} = q_0^{l-2}$. Since the map $X_l \to X_{l-1}$ is of degree q_0, we obtain $\deg(x_{l-1})_0^{(l)} = q_0 \deg(x_{l-1})_0^{(l-1)} = q_0^{l-1}$. Summarizing,

$$h = 2q_0^{l-1} = \frac{2n_l}{q_0^2 - 1}.$$

Now the parameters in (18) are obtained from (10) by direct computation. ■

2) : Now consider the second natural projection of curves in the tower (16). Namely, let Y_l correspond to the function field $\Bbbk(z_2, \ldots, z_l)$ and consider the field embedding

$$(g_l')^* : \Bbbk(Y_l) \to \Bbbk(X_l) = \Bbbk(x_1, z_2, \ldots, z_l).$$

Note that g_2' is the projection $g' : X_2 \to \mathbb{P}^1$ considered in Section IV-A2. The curves $\{Y_l, l = 2, 3, \ldots\}$ form another optimal tower of curves [2, Remark 3.11] given by the recursive equations

$$Y_l : z_l^q + z_l = \frac{z_{l-1}^q}{z_{l-1}^{q-1} + 1},\ l \geq 3;\ Y_2 := \mathbb{P}^1.$$

In geometric terms, the embedding $(g_l')^*$ implies that the curve X_l is the fiber product of X_2 and Y_l over $Y_2 = \mathbb{P}^1$, viz.

$X_l = X_2 \times_{Y_2} Y_l$, which in turn implies that the projection $g'_l : X_l \to Y_l$ shares the main properties of $g' = g'_2$. Indeed, we have:

1) The genus of Y_l satisfies $G'_l < q_0^{l-1}$ (the exact value is given in [2, Remark 3.8]; note that the notation for $\mathbb{k}(Y_l)$ in [2] is T_{l-1}).
2) Let $\pi_l : Y_l \to Y_2$ be the natural projection of degree $\deg(\pi_l) = q_0^{l-2}$. All the points in $S_l := \pi_l^{-1}(\mathbb{F}_q \backslash M)$ are $\mathbb{F}_q$-rational and
$$|S_l| = q_0^{l-2}(q_0^2 - q_0) = q_0^{l-1}(q_0 - 1) = n_l/(q_0 + 1).$$
3) The point $\infty \in Y_2 = \mathbb{P}^1$ is totally ramified, i.e., $\pi_l^{-1}(\infty) = P'_{\infty,l}$ for a rational point $P'_{\infty,l} \in Y_l$.
4) We have $(g'_l)^{-1}(S_l) = (\psi_l)^{-1}(\mathbb{F}_q), |(g'_l)^{-1}(S_l)| = n_l$, and all the points in $(g'_l)^{-1}(S_l)$ are $\mathbb{F}_q$-rational. The degree of the projection g'_l is $\deg(g'_l) = q_0 + 1$. The fibers of g'_l are transversal with those of g_l.
5) The degree of $x_1 : X_l \longrightarrow \mathbb{P}^1$ equals $h := \deg(x_1) = \deg(\pi_l)\deg((x_1)_0^{(2)}) = q_0^{l-1}$.

We obtain the following statement.

Proposition 4.5: There exists a family of q-ary $(n, k, r = q_0)$ LRC codes on the curve $X_l, l \geq 2$ with the parameters
$$\left.\begin{array}{c} n = n_l = q_0^{l-1}(q_0^2 - 1) \\ k \geq r(t - q_0^{l-1} + 1) \\ d \geq n_l - t(q_0 + 1) - (q_0 - 1)q_0^{l-1} \end{array}\right\} \qquad (19)$$
where t is any integer such that $G_{l-1} \leq t \leq n_{l-1}$.

Proof: Put $r = q_0, \ell = 1$ and apply the construction of Theorem 3.1 to
$$X := X_l, Y := Y_l, g := g'_l, Q_\infty := P'_{\infty,l}, D_t = tQ_\infty. \qquad \blacksquare$$

Remark 4.2: For the construction of Prop. 4.5 the lower bound of Remark 4.1 takes the form $h \geq n_l/q_0^2 = q_0^{l-1} - q_0^{l-3}$ which is very close the actual value $h = q_0^{l-1}$. In the case of Prop. 4.4 the value h is about twice as large as the lower bound.

Remark 4.3: Due to the results of [6], the basis of the function space $L(D_t)$ and the set S_l can be found in time polynomial in n_l, and so the codes of Prop. 4.5 are polynomially constructible.

C. Asymptotic bounds

Let us compute the asymptotic relation between the parameters of LRC codes on the Garcia-Stichtenoth curves constructed above.

Proposition 4.6: Let $q = q_0^2$, where q_0 is a prime power. There exist families of LRC codes with locality r whose rate and relative distance satisfy
$$R \geq \frac{r}{r+1}\Big(1 - \delta - \frac{3}{\sqrt{q}+1}\Big), \qquad r = \sqrt{q} - 1 \qquad (20)$$
$$R \geq \frac{r}{r+1}\Big(1 - \delta - \frac{2\sqrt{q}}{q-1}\Big), \qquad r = \sqrt{q}. \qquad (21)$$

Remark 4.4: Recall that without the locality constraint the relation between R and δ for codes on asymptotically optimal curves (for instance, on the curves $X_l, l = 2, 3, \dots$) takes the form $R \geq 1 - \delta - \frac{1}{\sqrt{q}-1}$, see [12, p.251].

Proof: For instance, let us check (20). From (18) we obtain
$$\begin{aligned} d + \frac{k(r+1)}{r} &\geq n_l - \frac{q_0 n_{l-1}}{q_0 - 1} - \frac{2n_l(q_0 - 2)}{q_0^2 - 1} + q_0 \\ &\geq n_l\Big(1 - \frac{1}{q_0 - 1} - \frac{2q_0 - 4}{q_0^2 - 1}\Big) \\ &= n_l\Big(1 - \frac{3}{q_0 + 1}\Big). \end{aligned}$$
Letting $\delta = d/n_l, R = k/n_l, l \to \infty$, we obtain (20). $\blacksquare$

One of the main results about the classic AG codes is an improvement of the Gilbert-Varshamov (GV) bound starting with $q = 49$. A GV-type bound for LRC codes was recently obtained in [13].

Theorem 4.7: There exists a sequence of q-ary linear r-LRC codes with the parameters (R, δ) as long as
$$R < \frac{r}{r+1} - \min_{0 < s \leq 1}\Big\{\frac{1}{r+1}\log_q b(s) - \delta \log_q s\Big\}, \qquad (22)$$
where
$$b(s) = \frac{1}{q}((1 + (q-1)s)^{r+1} + (q-1)(1-s)^{r+1}). \qquad (23)$$
The bound given in (21) (i.e., the code family constructed in Prop. 4.5) improves upon the GV bound for large alphabets. For instance, for $q_0 = 23$ the code rate (21) is better than (22) for $\delta \in [0.413, 0.711]$, and the length of this interval increases as $q_0 \to \infty$. Similar conclusions can be made for the codes in the family of Prop. 4.4.

V. Concluding remarks

Note that the locality parameter r for q-ary LRC codes obtained from the Garcia-Stichtenoth curves is fixed and is equal to about $\sqrt{q}$. Generally one would prefer to construct LRC codes for any given r, or at least for a range of its values. It is possible to modify the construction of this paper to attain small locality (such as, for instance, $r = 2$), while still obtaining code families that improve upon the GV bound (22). This construction will be presented in a longer version of this extended abstract [14].

References

[1] A. Garcia and H. Stichtenoth, *A tower of Artin-Schreier extensions of function fields attaining the Drinfeld-Vladut bound*, Invent. math. **121** (1995), 211–222.

[2] ———, *On the asmptotic behaviour of some towers of functions fields over finite fields*, J. Number Theory **61** (1996), no. 2, 248–273.

[3] P. Gopalan, C. Huang, H. Simitci, and S. Yekhanin, *On the locality of codeword symbols*, IEEE Trans. Inform. Theory **58** (2011), no. 11, 6925–6934.

[4] D. S. Papailiopoulos and A. G. Dimakis, *Locally repairable codes*, Proc. 2012 IEEE Internat. Sympos. Inform. Theory, 2012, pp. 2771–2775.

[5] A.S. Rawat, D.S. Papailiopoulos, A.G. Dimakis, and S. Vishwanath, *Locality and availability in distributed storage*, Proc. 2014 IEEE Int. Sympos. Inform. Theory (Honolulu, HI), pp. 681–685.

[6] K. W. Shum, I. Aleshnikov, P.V. Kumar, H. Stichtenoth, and V. A. Deolalikar, *A low-complexity algorithm for the construction of algebraic-geometric codes better than the Gilbert-Varshamov bound*, IEEE Trans. Inform. Theory **47** (2001), no. 6, 2225—2241.

[7] N. Silberstein, A. S. Rawat, O. O. Koyluoglu, and S. Vishwanath, *Optimal locally repairable codes via rank-metric codes*, arXiv:1301.6331.

[8] I Tamo and A. Barg, *Bounds on locally recoverable codes with multiple recovering sets*, Proc. 2014 IEEE Int. Sympos. Inform. Theory (Honolulu, HI), pp. 691–695.

[9] I. Tamo and A. Barg, *A family of optimal locally recoverable codes*, IEEE Trans. Inform. Theory **60** (2014), no. 8, 4661–4676.

[10] I. Tamo, A. Barg, S. Goparaju, and A. R. Calderbank, *Cyclic LRC codes and their subfield subcodes*, arXiv.org, 2015.
[11] I. Tamo, D. S. Papailiopoulos, and A. G. Dimakis, *Optimal locally repairable codes and connections to matroid theory*, Proc. 2013 IEEE Internat. Sympos. Inform. Theory, 2013, pp. 1814–1818.
[12] M. Tsfasman, S. Vlăduţ, and D. Nogin, *Algebraic geometric codes: Basic notions*, Mathematical Surveys and Monographs, vol. 139, American Mathematical Society, Providence, RI, 2007.
[13] A. Barg, A. Frolov, and I. Tamo, *Bounds on the parameters of locally recoverable codes*, 2015, in preparation.
[14] A. Barg, I. Tamo, and S. Vlăduţ, *LRC codes on algrebraic curves*, in preparation.